

Genforce version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 5, 2001, 14:29:18 ; Search time 58.44 Seconds
(without alignments)
216.295 Million cell updates/sec

Title: US-09-593-316-2

Sequence: 1 MNVKGKYLMLVYVIVV.....IKLVKMSQTKRVNVVNNV 369

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	1906.5	95.2	368 1	GATR_BOVIN
2	1711	85.4	371 1	GATR_PIG
3	1471	73.4	394 1	GATR_MOUSE
4	670	33.4	354 1	BOAT_HUMAN
5	110	5.5	505 1	AGAR_ALTAT
6	96	4.8	583 1	PR40_YEAST
7	94.5	4.7	408 1	Y233_METJA
8	94.5	4.7	563 1	YKNO_YEAST
9	94.5	4.7	688 1	TALA_POVIC
10	93	4.6	1057 1	POL_STVAT
11	91.5	4.6	522 1	HEX1_ENTHI
12	89.5	4.5	959 1	DPO1_AERPE
13	89	4.4	699 1	HS82_ORYSA
14	88.5	4.4	517 1	DMEN_PESP
15	87	4.3	536 1	VHT1_HRT7
16	87	4.3	569 1	GCL1_DROME
17	86.5	4.3	510 1	HVAL_MACEA
18	86.5	4.3	691 1	TALA_POVHA
19	86.5	4.3	733 1	YLC4_CAFEL
20	86.5	4.3	2410 1	POLL_BAYMJ
21	85.5	4.3	708 1	TALA_SVA0
22	85	4.2	863 1	PHSC_MYCTU
23	85	4.2	1276 1	BXD_CIOBO
24	84.5	4.2	185 1	YQIM_YHAST
25	84.5	4.2	514 1	THRC_SCHPO
26	84.5	4.2	558 1	OL62_BACCE
27	84.5	4.2	2386 1	RAD3_SCHPO
28	84	4.2	408 1	ARCA_MYCHO
29	84	4.2	446 1	GLNA_MCTIP
30	84	4.2	727 1	IF2M_BOVIN
31	84	4.2	1134 1	YML7_YEAST
32	83.5	4.2	434 1	KP58_MOUSE
33	83.5	4.2	461 1	SSL1_YEAST

34	83.5	4.2	467 1	IRF6_HUMAN	014896 homo sapien
35	83.5	4.2	502 1	MTF6_AZOVI	P11067 azotobacter
36	83	4.1	353 1	Y951_METJA	Q58361 methanococ
37	83	4.1	639 1	GLOB_HUTFI	P10539 butyivibrio
38	83	4.1	992 1	VP41_YEAST	P148959 saccharomy
39	83	4.1	1451 1	SP16_YEAST	P23615 saccharomy
40	82.5	4.1	368 1	MNN9_CANAL	P34697 candida alb
41	82.5	4.1	436 1	KP58_HUMAN	P21127 homo sapien
42	82.5	4.1	1586 1	SNZ2_HUMAN	P15131 homo sapien
43	82	4.1	454 1	GLNA_METJA	Q60182 methanococ
44	82	4.1	559 1	PAG1_BOVIN	Q07547 bos taurus
45	82	4.1	699 1	HS80_LYCES	P16181 lycopersico

ALIGNMENTS

```

RESULT 1
GATR_BOVIN STANDARD: PRT: 368 AA.
ID GATR_BOVIN
AC P14769:
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 30-MAY-2000 (Rel. 39, Last annotation update)
DE N-ACETYLGLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE (EC 2.4.1.151)
DE (GALACTOSYLTRANSFERASE) (UDP-GALACTOSE: BETA-D-GALACTOSYL-1,4-N-ACETYL-
DE D-GLUCOSAMINIDE ALPHA-1,3-GALACTOSYLTRANSFERASE).
GN GTRAT.
OS Bos taurus (bovine).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovidae; Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE:89340543; PubMed:2503516;
RA Joziasse D.H., Shaper J.H., van den Eijnden D.H., van Tienen A.J.,
RA Shaper N.L.:
RT "Bovine alpha 1->3-galactosyltransferase: isolation and
RT characterization of a cDNA clone. Identification of homologous
RT sequences in human genomic DNA."
RL J. Biol. Chem. 264:14290-14297(1989).
CC -!- FUNCTION: TRANSFER OF GALACTOSE FROM UDP-GALACTOSE TO AN
CC ACCEPTOR MOLECULE (R).
CC -!- CATALYTIC ACTIVITY: UDP-GALACTOSE + BETA-D-GALACTOSYL-1,4-
CC N-ACETYL-D-GLUCOSAMINYL-R = UDP + ALPHA-D-GALACTOSYL-1,3-
CC BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINYL-R.
CC -!- COFACTOR: MANGANESE.
CC -!- PATHWAY: GLYCOSYLATION.
CC -!- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS CISTERNAE OF GOLGI.
CC -!- DISEASE: AUTOIMMUNE DISEASE (ANTI-BODIES AGAINST ENZYMAIC
CC PRODUCT).
CC -!- SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERASE.
CC -!- SIMILARITY: STRUCTURAL SIMILARITY WITH THE OTHER MAMMALIAN
CC GLYCOSYLTRANSFERASES.
-----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: J04989; AAA30558.1; -.
DR PIR: A44785; A44785.
KW Signalase: Glycosyltransferase; Glycoprotein; Transmembrane;
KW Signal-anchor; Golgi stack; Manganese.
KW DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
KW SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).
FT DOMAIN 7 22 LUMENAL, CATALYTIC (POTENTIAL).
FT DOMAIN 23 368 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHYD 293 293 N-LINKED (GLCNAG, ...) (POTENTIAL).

```


CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90046769; PubMed=2510162;
RA Larsen R.D., Rajan V.P., Kuff M.M., Kukowska-Latallo J.,
RA Cummings R.D., Lowe J.B.;
RT "Isolation of a cDNA encoding a murine
KT UDPgalactose-4-epimerase (EC 5.1.3.1) from mouse liver."
R1 alpha-1,3-galactosyltransferase: expression cloning by gene
R1 transfer.";
RL Proc. Natl. Acad. Sci. U.S.A. 86:8227-8231(1989).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=92184813; PubMed=1544928;
RA Joiasse D.H., Shaper N.L., Kim D., van den Eljnden D.H., Shaper J.H.;
RT "Murine alpha 1,3-galactosyltransferase. A single gene locus
R1 specifies four isoforms of the enzyme by alternative splicing.";
RL J. Biol. Chem. 267:5534-5541(1992).
CC -1- FUNCTION: TRANSFER OF GALACTOSE FROM UDP-GALACTOSE TO AN
CC ACCEPTOR MOLECULE (R).
CC -1- CATALYTIC ACTIVITY: UDP-GALACTOSE + BETA-D-GALACTOSYL-1,4-
CC ACETATEL-1,3-GLUCOSAMINYL-R = UDP + ALPHA-D-GALACTOSYL-1,3-
CC BETA-D-GALACTOSYL-1,4-N-ACETYL-D-GLUCOSAMINYL-R.
CC -1- COFACTOR: MANGANESE.
CC -1- PATHWAY: GLYCOSYLATION
CC -1- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN. MEMBRANE-BOUND
CC FORM IN TRANS OLIGOMER OF GOLGI.
CC -1- SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERRIN.
CC -1- SIMILARITY: STRONG, TO BLOOD GROUP AB TRANSFERRIN
CC GLYCOSYLTRANSFERASES.
CC
CC
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC
CC
DR EMBL; M26925; AAA37657.1; -;
DR EMBL; M85153; AAA37711.1; -;
DR PIR; A34417; A34417.
DR MGD; MG1:95704; Galat.
KW Transferase; Glycosyltransferase; Glycoprotein; Transmembrane;
KM Signal-anchor; Golgi stack; Alternative splicing; Manganese.
FT DOMAIN 1 41
FT TRANSMEM 42 60
FT DOMAIN 61 394
FT CARBOHD 83 83 LUMENAL, CATALYTIC (POTENTIAL).
FT CARBOHD 319 319 N-LINKED GLUCAN. . .) (POTENTIAL).
FT VARSPPLIC 1 35 MISSING (IN SHORTER ISOFORM).
FT VARSPPLIC 62 62 MISSING (IN SHORTER ISOFORM).
SQ SEQUENCE 394 AA: 46475 MW: 7766816DaOHHBF GCRG64;

[illegible]

CY	178	ELGSLRSFFKFEVKKPERRRWQDVSMMVKMTIGEHVAHIOREDFLFCMOVDGVFODPHEGV	234
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
D6	203	HLLNFLLSVFLEIRSEPKRWDISMMRKKTITGEHLAHIOHFVDPLFCMVAOVVOVDNGV	262
CY	238	ETLLESYAQLQAWMYADPEEFYEERKESSAVIPFGEDHPFYHAALPGSPPTVLNITO	297
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
D6	263	EFTLSQLOLVLAQWAMYKASPEKFTFERRELISAATIPGRSDPYHAALPGSPPTHLNLTR	322
CY	298	ECEFGILLKDKKNDIENAGWHDSHLNKYPFLILNFKTLLSPEYWDHYGLPADIKLVKSM	357
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
D6	323	ECFGIILDDCKHDLEQMHDESHLNKFLENRPTKILTSPEYCMVDQICLPSIDSKSVKVM	382
CY	358	OTKEYNVKNNV	369
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
D6	383	OTKEYNLVRNNV	394
		: : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :	
RESULT	4		
MGAT_HUMAN			
ID	BGAT_HUMAN	STANDARD:	PRT: 354 AA.
AC	P16442:		
DT	01-AUG-1990	(Rel. 15, Created)	
DT	01-DEC-1992	(Rel. 24, Last sequence update)	
DT	01-OCT-2000	(Rel. 40, Last annotation update)	
DE	HISTO-BLOOD GROUP ABO SYSTEM TRANSFERASE (MGAT) [INCLUDES: GLYCOPROTEIN-FUCOSYLTRANSFERASE SIDE ALPHA-N- ACTEYL-GALACTOSAMINYLTANSFERASE (EC 2.4.1.40) GLYCOPROTEIN-FUCOSYLTRANSFERASE (EC 2.4.1.40) ALPHA-N-ACEYL-LACTOSAMINYLTANSFERASE) (HISTO-BLOOD GROUP A TRANSFERASE) (A TRANSFERASE): GLYCOPROTEIN-FUCOSYLGALACTOSE ALPHA- GALACTOSULTANSFERASE (EC 2.4.1.37) (FUCOSYLGALCOPROTEIN + ALPHA- GALACTOSULTANSFERASE) (HISTO-BLOOD GROUP B TRANSFERASE) (B TRANSFERASE)]].		
DN	ABO.		
GN	Homo sapiens (Human).		
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;		
OC	Eumariota; Euteleostomi; Primates; Catarrhini; Hominoidea; Homo.		
OX	NCB1:taxid=9606:		
RN	[1]		
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.		
RX	MEDLINE=90110098: PubMed=2104828:		
RA	Yamamoto F.-I., Marken J., Tsuji T., White T., Clausen H.,		
RL	Hakomori S.-I.;		
RT	"Cloning and characterization of DNA complementary to human UDP-		
RJ	GalNAc: Fuc alpha 1->2Gal alpha 1->3GalNAc transferase (histo-blood		
RL	group A transferase) mRNA."		
RN	J. Biol. Chem. 265:1146-1151(1990).		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=90238543: PubMed=2333095:		
RA	Yamamoto F.-I., Clausen H., White T., Marken J., Hakomori S.-I.;		
RL	"Molecular genetic basis of the histo-blood group ABO system."		
RJ	Nature 345:229-233(1990).		
RN	[3]		
RP	CHARACTERIZATION,		
RX	MEDLINE=91035461: PubMed=2121736:		
RA	Yamamoto F.-I., Hakomori S.-I.;		
RL	"Sugar-mucopolide donor specificity of histo-blood group A and B		
RJ	transferases is based on amino acid substitutions."		
RL	J. Biol. Chem. 265:19257-19262(1990).		
CC	-1- FUNCTION: THIS PROTEIN IS THE BASIS OF THE ABO BLOOD GROUP SYSTEM.		
CC	THE HISTO-BLOOD GROUP ABO INVOLVES THREE CARBOHYDRATE ANTIGENS: A,		
CC	B, AND H. A, B, AND AB INDIVIDUALS EXPRESS A GLYCOSYLTRANSFERASE		
CC	ACTIVITY THAT CONVERTS THE H ANTIGEN TO THE A ANTIGEN (BY ADDITION		
CC	OF UDP-GALNAc) OR TO THE B ANTIGEN (BY ADDITION OF UDP-GAL).		
CC	WHEREAS O INDIVIDUALS LACK SUCH ACTIVITY.		
CC	-1- CATALYTIC ACTIVITY: UDP-N-ACETYL-D-GALACTOSAMINE + GLYCOPROTEIN		
CC	ALPHA-1-FUCOSYL-(1,2)-D-GALACTOSE -> UDP + N-ACETYL-ALPHA-D-		
CC	GALACTOSAMINYL-(1,3)-[ALPHA-1-FUCOSYL-(1,2)]-D-GALACTOSE.		
CC	-1- CATALYTIC ACTIVITY: UDP-GALACTOSE + GLYCOPROTEIN ALPHA-L-FUCOSYL-		
CC	(1,2)-D-GALACTOSE -> UDP + GLYCOPROTEIN ALPHA-D-GALACTOSYL-(1,4)-		
CC	[ALPHA-L-FUCOSYL-(1,2)] D-GALACTOSE.		
CC	-1- PATHWAY: GLYCOSYLATION.		
CC	-1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN; MEMBRANE-BOUND		

[illegible]

```

07 345 SPEYGVNTH IGLFDIKLKXKSWQTEKEVNVN 367
    ||||| 11 -||||| : : : : : 11
DB 420 SPETIMDQILLGMIAVLKRLEITAVTKRHOAVRN 453

RESULT 5
AGAR ALTAT
ID AGAR ALTAT STANDARD; PROT: 50% AA.
AC P13734;
DE 01-JAN-1990 (rel. 13, created)
DE 01-JAN-1990 (rel. 14, last sequence update)
DE 15-JUL-1999 (rel. 38, last annotation update)
DE BETA-AGARASE PRECURSOR (EC 3.2.1.81).
GN AGAR.
OS Achromonas atlantica (Pseudomonas atlantica);
OC Bacteria; Proteobacteria; gamma subdivision; Achromonadaceae;
CC Pseudomonadotomones;
CX NCBI_TaxID=288;
RN SEQUENCE FROM N.A.
RP MEDLINE 89121082; PUBMED-2914859;
RA Helms R.;
RT "Sequence analysis of the agar gene encoding beta agarase from
   Pseudomonas atlantica."
RL J. Bacteriol. 171:602-605(1989).
CC 1- CATALYTIC ACTIVITY: HYDROLYSIS OF 1,3-BETA-D-GALACTOSIDE LINKAGES
   IN AGAROSE, GIVING THE TETRAMER AS THE PREDOMINANT PRODUCT.
CC 2- SUPERCELLULAR LOCATION: SECRETED.
CC 3- DATABASE: NAME-Worthington biochem.com/manual/A/Agar.html";
CC WWW="http://www.worthington-biochem.com/manual/A/Agar.html".
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL institution.
CC The European Bioinformatics Institute. There are no restrictions on ways
CC use by non-profit institutions as long as their content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sb.ch).
-----
DB PMBL: M22725; AAA25696.1;
DB PIR: A32261; A32261.
KM Hydrolyase; Glycosidase; Signal.
PT SIGNAL 1 24
FT CHAIN 24 505 BETA-AGARASE.
FT ACT SITE 200 200 PROTON DONOR (POTENTIAL).
FT ACT SITE 322 322 NOCTEPHILE (POTENTIAL).
SO SEQUENCE 505 AA; 57617 MW; 66B75CBF486280B1 CRC64;

Query Match 5.58; Score 110; DB 1; Length 505;
Best Local Similarity 19.08; Pred. No. 0.15;
Matches 82; Conservative 53; Mismatches 145; Indels 154; Gaps 17.

07 6 KVLMSLVVS-----TVIVFWMEYHSIPGSLFWINSPNPVSGSSSTQKQM 54
    ||| ||| | : ||| ||| ||| : : : :
DB 3 KVLFWLLVTSLSVAIPPIYIATTEVVVNLNVRKSYSVGKSEF---EKKNNILKHSTINDM 59

07 54 WPFWMINNGYOEDENVDERE-----QRKESKLKLSIMWFNPKRPVVTMTW 104
    || : : : : : || : : : : : || :
DB 60 WQGFEEKLRKVMFEELDYFCRGNGGTVMWNQAIEDPANIGADPNIIIAKGCAQRETNW 119

07 105 KAIVWEGHTNRKAVLDIYAKKQITV-----LTAVAGRYIE 142
    || : : : : : || : : : : : || :
DB 120 -----GQNKSAHQHCYCKGHMLMGCGPRAYLCNISPGCVGSAMQAKGAAGVGPFLG 171

07 143 HYLFEELTSANKPMCHRVIFPVMMVIDVSRMDL-----ELG-----PLR 183
    || : : || : : || : : : : || : :
DB 172 QYNVEFFRSACDVFYKSHLPYE--EVLNHPYGVTTAPNHDLKPDPPIIDITFN 228

07 184 -----SFVFEVKPKRKQDVSMVRKKTGGHIYVHGIGREVDLF 223
    : : || : : || : : : : : :
DB 229 IVADAFGQHNTHTKLGTFVAAPFLFEQRFVFAWFEREKLEIDTSGSHIMVYSTHYD---285

```

QY 224 CNDVDFE-----DEGVETLGESE-----AQLQAWW- 252
 DB 286 -LEDDNREFKSRLEATLDMIDVSLALGETKPHVISEYSGHRRPENAIPWALRKDWPL 344
 QY 253 -KAUPEFYERKREBSA-----ATIPGEGDFEYHAAIFG-----GTFYQVNLITQEGFKCI 303
 DB 345 KIASPMLMQPLSPDVSILSTIPVPI-----KALMGTADQGTYYNKRLLRQV----- 391
 QY 304 LKDKNDIEAQM 315
 DB 392 -KEAPNTEGENW 402

RESULT 6
 PR40_YEAST STANDARD; PRT: 583 AA.
 ID PR40_YEAST STANDARD; PRT: 583 AA.
 AC P33203;
 DT 01-FEB-1994 (rel. 28, last sequence update)
 DT 01-FEB-1994 (rel. 28, last sequence update)
 DT 01-OCT-2000 (rel. 40, last annotation update)
 DE PRE-MRNA PROCESSING PROTEIN PRP40.
 GN PRP40 OR YKL012W OR YKL165.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN 111
 RP SEQUENCE FROM N.A.
 RX MEDLINE=93127732; PubMed=1481574;
 RA Pascolo S., Ghazvini M., Boyer J., Colliaux L., Thierry A., Dujon B.;
 RT "The sequence of a 9.3 kb segment located on the left arm of the
 RT yeast chromosome XI reveals five open reading frames including the
 RT CCE1 gene and putative products related to MYO2 and to the ribosomal
 RT protein L10.*"
 RL Yeast 8:987-995(1992).
 RN 121
 RP IDENTIFICATION
 RX MEDLINE=96182100; PubMed=8622699;
 RA Kao H.-Y., Siliciano P.G.;
 RT "Identification of Prp40, a novel essential yeast splicing factor
 RT associated with the U1 small nuclear ribonucleoprotein particle.*"
 RL Mol. Cell. Biol. 16:960-967(1996).
 RN 131
 RP PARTIAL SEQUENCE.
 RX MEDLINE=97165029; PubMed=9012791;
 RA Neubauer G., Gottschalk A., Fabrizio P., Seraphin B., Luehrmann R.,
 RA Mann M.;
 RT "Identification of the proteins of the yeast U1 small nuclear
 RT ribonucleoprotein complex by mass spectrometry.*"
 RL Proc. Natl. Acad. Sci. U.S.A. 94:385-390(1997).
 CC -1- FUNCTION: REQUIRED FOR THE FIRST STEP OF PRE-MRNA SPLICING. THIS
 CC PROTEIN IS ASSOCIATED WITH SNRP U1.
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.
 CC -1- SIMILARITY: CONTAINS 2 WW DOMAINS.
 CC -1- SIMILARITY: WEAK SIMILARITY TO YEAST MYO2.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL: S53418; AAB24902.1; -
 DR EMBL: Z28012; CAAB1847.1; -
 DR PIR: S30014; S30014.
 DR SDD: S0001495; PRP40.
 DR InterPro: IPR002713; -
 DR InterPro: IPR002713; -
 DR Pfam: PF01846; FF: 4.
 DR Pfam: PF00397; WW: 2.

DR PROSITE: PS01159; WW DOMAIN_1; 2.
 DR PROSITE: PS00020; WW DOMAIN_2; 2.
 KM mRNA processing; mRNA splicing; Nuclear protein; Ribonucleoprotein;
 KW Repeat.
 FT DOMAIN 1 31 WW.
 FT DOMAIN 42 72 WW.
 FT MUTAGEN 240 240 S->P: IN PRP40-1 SUPPRESSOR AFFECTS SAK1
 FT MRNA ACCUMULATION IN U1-04 MUTANT AT 18
 FT DEGREES CELSIUS.
 SQ SEQUENCE 583 AA: 69065 MW: BAICIGYIDJ532524C CCG64;

Query Match 4.8%; Score 96; Dh 1; length 583;
 Best Local Similarity 20.9%; Pred. No. 2.5;
 Matches 67; Conservative 51; Mismatches 101; Indels 102; Gaps 20;

QY 67 DEDVEEKKRREKRSKLT--KLSDWENPKRPEVYMTLWKAFAVVECTYNNKAVLDYYA 124
 DB 256 DILDTQKSKKKKLTQALKELREYLN-----GIITSSSEFTITWQQLNHYVD---- 306
 QY 125 KQKITVGLTFVAVGRIEYEEFLTSANKRPV-GHRYT---FVYVDDVSRLMLELG 180
 DB 307 KSK-----RYM-----ANRHPKVLTHEGVNLEYLKIYNTIENDLQNTLN 345
 QY 181 PIR-----SFK-----VFVYKPRRQDVSVKRTTIGENIYVATITQREVDPL 222
 DB 346 ELRLRYTDRARDNFKSLREVPPIKANTKMSD-----IYPIKSDPRFL 393
 QY 223 F-----CNDVDFE-----FGVETLGESEVAQLQAW-WYKADPEFYERKRESA 268
 DB 394 HMLGRGSSCLDLPLDFVEQDMYIFAQKSLAQGLTLIDQNFPMNADISDEITKQNEVL 453
 QY 269 ATIPGEGFYHAAIFGTFYQVNLITQEGFKGILKDKKNDIPKQMDPSL-----NKY 324
 DB 454 -----END-----RKFDKVDKEDISLVD---GLIKQENKIDQKLNREKLEDKKHY 499
 QY 325 F--LNKP--TKLSPK--YQMD 341
 DB 500 FWLLDRTYTKIGKRPSTWD 520

RESULT 7
 Y233_METJA STANDARD; PRT: 408 AA.
 ID Y233_METJA STANDARD; PRT: 408 AA.
 AC Q60290;
 DT 01-NOV-1997 (rel. 35, last sequence update)
 DT 01-NOV-1997 (rel. 35, last sequence update)
 DT 01-NOV-1997 (rel. 35, last sequence update)
 DE HYPOTHETICAL PROTEIN MJECL33.
 GN MJECL33.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
 OC Methanococcus.
 OX NCBI_TaxID=2190.
 RN 111
 RP SEQUENCE FROM N.A.
 RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bull C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., Fitzgerald J.M., Clayton R.A., Gocayne J.D.,
 RA Kerlavage A.R., Dougherty B.A., Tomb L.F., Adams M.D., Keich C.L.,
 RA Overbeek R., Kirkness E.P., Weinstock K.G., Merrick J.W., Glodek A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Ulteback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Colton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klotz H.-P., Fraser C.M., Smith H.O., Moese G.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.*"
 RL Science 273:1058-1073(1996).
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its

KL Plantia 193:57-66(1994).
 CC -1- FUNCTION: MOLECULAR CHAPERONE. HAS ATPASE ACTIVITY
 CC (BY SIMILARITY).
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC (POTENTIAL).
 CC -1- SIMILARITY: BELONGS TO THE HEAT SHOCK PROTEIN 90 FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.sib.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 DR EMBL: Z11920; CAA77978.1; -
 DR PIR: S25541; S25541.
 DR HSSP: P07900; 1YER.
 DR InterPro: IPR001404; -
 DR Pfam: PF00183; HSP90.1.
 DR PRINTS: PR00775; HEATSHOCK90.
 DR PROSITE: PS00298; HSP90.1.
 KM Chaperone; ATP-binding; Heat shock.
 SO SEQUENCE 699 AA: 80194 MW: 538104572838CE CRC64;

 Query Match 4.4%; Score 89; DB 1; Length 699;
 Best Local Similarity 20.4%; Pred. No. 12;
 Matches 64; Conservative 49; Mismatches 98; Indels 102; Gaps 17;

 QY 23 EYHSPEGLFPMINPSKNPVSGSSIOKGMWPFKMFNNNOYOEDEDVYHKKQREKDS 82
 DB 201 EPISTYPS-IMIEKTEKTSIDDEDEEK-----KDEBKRVKVDVDEKKEKKKK 251
 QY 83 KLK-LSDF-----NP-FKREVVMTDMKAPVWEGTYNRAVLDDYAKOKITVGLT 133
 DB 252 KIKVSHHSLVANKKIMKRPKEITKEVAA-----FYSLINDW----- 293
 QY 134 VPAVKRIEYHLEFITSANKHGVGHVYVYVVDVSKMPL-----TELGL 182
 DB 294 -----EEHL-AVKHESVQLFEKAVLLEVPKRAPDLEDTKKLNKLYVR 339
 QY 183 KSP-----KVEFVPRKRWODVSWKMTIGEHIVAHIORE-----VD 220
 DB 340 KVTIMNGCELLE-W--LSFYKQIVDSHDLPLNISRELGQNKILKLYRKLVKVCVE 395
 QY 221 ELFCM-----DVDVFD-----EFVY---ETLGEVAVLOAMVYADPDETYERKES 267
 DB 396 LFEFIANKKDYKHYKAFSKNLKLGTHHSTNRNKTAHLRYHSTKSGDEL--SLKDY 453
 QY 268 AAVIPFEGCGPY 280
 DB 454 VTRMKRQNDIY 466

 RESULT 14
 DMPN_PESP STANDARD; PRT; 517 AA.
 ID DMPN_PESP
 AC P19732;
 DT 01-FEB-1991 (Rel. 17, Created)
 DT 01-FEB-1991 (Rel. 17, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE PHENOL HYDROXYLASE P3 PROTEIN (PC 1.14.13.7) (PHENOL 2-MONOXYGENASE
 DE P3 COMPONENT).
 GN DMPN OR PHEA4.
 OS Pseudomonas sp. (strain CF600).
 OG Plasmid pV1150.
 OC Bacteria; Proteobacteria.
 OX NCBI_LtaxID=106;
 RN 11;
 RP SEQUENCE FROM N.A.
 RX MEDLINE=91072230; PubMed=2254258;
 RA Nordlund I., Powlowski J., Shingler V.;
 KT *Complete nucleotide sequence and polypeptide analysis of

KT multicomponent phenol hydroxylase from pseudomonas sp. strain
 RT CF600.?
 KL J. Bacteriol. 172:6826-6833(1990).
 RN 12;
 RP SEQUENCE FROM N.A.
 RC STRAIN-BH;
 RA Takeo M., Maeda Y., Okada H., Miyama K., Mori K., Ike M.,
 RA Fujita M.;
 KL Submitted (MAR-1994) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: CATABOLIZES PHENOL, AND SOME OF ITS METHYLATED
 CC DERIVATIVES. P3 IS REQUIRED FOR GROWTH ON PHENOL, AND FOR
 CC IN VITRO PHENOL HYDROXYLASE ACTIVITY.
 CC -1- CATALYTIC ACTIVITY: PHENOL + NADPH + O(2) -> CATECHOL + NADP(+) + H(2)O.
 CC -1- COFACTOR: FAD, AND REQUIRES FE(2+) FOR ACTIVITY.
 CC -1- PATHWAY: FIRST STEP OF PHENOL BIODEGRADATION.
 CC -1- SUBUNIT: THE MULTICOMPONENT ENZYME PHENOL HYDROXYLASE IS FORMED
 CC BY P0, P1, P2, P3, P4 AND P5 POLYPEPTIDS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb.sib.ch/announce/>
 CC or send an email to license@isb.sib.ch).
 CC -----
 DR EMBL: M60276; AAA25942.1; -
 DR PIR: D28864; BAA06017.1; -
 DR PIR: D37831; D37831.
 KM Aromatic hydrocarbons catabolism; Oxidoreductase; Monooxygenase;
 KM NADP; Flavoprotein; FAD; Iron; Plasmid.
 SO SEQUENCE 517 AA: 60522 MW: 46FD4B40ED74F9C CRC64;

 Query Match 4.4%; Score 88.5; DB 1; Length 517;
 Best Local Similarity 18.9%; Pred. No. 8.6; Indels 84; Gaps 12;
 Matches 59; Conservative 42; Mismatches 128;

 QY 72 EKKQREKDSKSLSDMNPFRKREVVMTDMKAPVWEGTYNRAVLDDYAKO----- 126
 DB 25 FITYQCKED-----VFPLFHHFGIKITDMK-----WEDFP-KITMTYKQYAKKK 71
 QY 127 KITVGLTFAVGRVIRHYIEEFLITSANKHGVGHVYVYVVDVSKMPLLEGLRSEK 186
 DB 72 KYAIFDAFAQNNQHNQISDARYNALKFLTAVSPLEVOAFQGESRVG-----RQFS 124
 QY 187 VFEVKKPRRWODVSWKMTIGEHIVAHIOREVDPL--FCMVVDVPODPHVEVETLGSNV 244
 DB 125 GAGARVACQMAIDELRHVQTVVHANSYKHHFDGLHDAIMYDKV----- 170
 QY 245 AOLQAMVYKADPD-----EFTYERKRESAAYIFPGCGPY-----YHAA 283
 DB 171 -----WYLSVRSKSYMDADRTAGPFEFLTAVSPSEFYVLTNLFVPRMGAAYNDDMATV 224
 QY 284 IFC-----GIPTVOLNITOCFKGLDKKKNDE--AQMHDSHLNKYPILN----- 328
 DB 225 TGFESAQSDSEARHMTLGLVEIKFMLEOHEDNVPITQKWDKFWKRYRLLITIGMMQY 284
 QY 329 KPTKILSPYCW 340
 DB 285 LPPKVMSSBAM 296

 RESULT 15
 VHTJ_BPT7 STANDARD; PRT; 536 AA.
 ID VHTJ_BPT7
 AC P03728;
 DT 21-JUL-1986 (Rel. 01, Created)
 DT 21-JUL-1986 (Rel. 01, Last sequence update)
 DT 01-FEB-1991 (Rel. 17, Last annotation update)
 DE HEAD-TO-TAIL JOINING PROTEIN.
 GN 8.

